

FIGURE 5

```

1  ATGGAACCTG AGTGCCTCGG ACATCACCCG GGTCTCATAA CTCTACGGCT GCAGCCCAAG
   w n l s a s d i t r v l k l y g c s p

61  TGGCCCCAGG CCCCCTGGGA GAGGGTCCCA TGCCACACAG ACTGGTAGGA GCCCCGCTCC
   s g p r p r g r g s h a h s t g r s p a

121 GGCCTCCCTA TCTCTGCAGC GGCTTTTGGA GGCACGTGTC GCGGAATCCA GGAGCCCCGA
   p a s l s l q r l l e a l s a e s r s p

181 CCCCAGTGGT TCCAGTGCGG GAGGCCAGCC CGTTCCTGCA GGGCCTGGGG AGAGCCCACA
   d p s g s s a g g q p v p a g p g e s p

241 TGGGTGGGAG TCCCCTGCCC TGAAAAAGCT CAGTGCAGAG GCCTCGGCAA GGCAGCCTCA
   h g w e s p a l k k l s a e a s a r q p

301 GACCCTAGCT TCCTCCCCAA GATCAAGGCC TGGAGCAGGT GCCCCCGGTG TTGCTCAGGA
   q t l a s s p r s r p g a g a p g v a q

361 GCAGTCCTGG CTGGCCGGAG TGTCCACCAA GCCACAGTC CCATCTTCAG AAGCAGGAAT
   e q s w l a g v s t k p t v p s s e a g

421 CCAGCCAGTC CCTGTCCAGG GAAGCCCAGC TCTGCCAGGG GGCTGTGTAC CTAGAAATCA
   i q p v p v q g s p a l p g g c v p r n

481 TTTCAAGGGG ATGTCCGAAG ATTAAGCCTG TGGCTTCTGT CCCCAGTAG GGAGGGCATC
   h f k g m s e d

541 CTCTGCCCAG TGGAGCTGGG TCGTCTACCT CTGGCTCCT TTGGGCCACA CCACTGTCTT
601 CCAGCCCCAA CCTACCACCC CATCTCAGAG GGCCAGGACT CTTCCCCTGT CTCTCTTCAC
661 TGTGTTCCCC TAAGGGCTCC TAGGGCCAGG GGTTCTTCTA GCTCTGCCAC AGGGGAAGGC
721 AGGCCTGGCT GTGCCTGCTC TTGACTTTTG CCCAGCCCTG GTGGATGCTG GGAATGGGAG
781 GTGACATTCT CCAGGGACAG GTCCTGGAAG GGGTGGGGAA GAGGTAGGTT CCAGCCCCGC
841 AGAACCCTGG AATCCCTCCT GTGCCTGAGG CCCTGCCCCC CAGCATGGAC TAATGGTGTC
901 CCTACCTCTC CCTCAGGGCA GCCCTGTGGC TGGGACCCTG GGAACAGCCT CCCATCCCAC
961 CCAACATGCC CAAGTGTGGG GGAATGTTCT ACAGCAGTGT AGCCTCCAGC CCTTCTCTCC
1021 AGGAGGCTTT GAGAGCCCAA CTTACTCCCC TGCAGAGCAG GAAGGTGGTA GGTCAGGTGT
1081 GGCCACCATT GGGGAGACGA GAAAGAAGTG GGGCCCCACC AGATTGCACA ATGGGAACCT
1141 CAGCTGGCCC CTGAACAGAG GACTCAGTTG TCTCCACCCT ACACCGCTAT TCCCTGGAGC
1201 TCAGCCAGGC GCAGCCTTGG AAGGAGAAAG GGCTGGGGTT ACCTGGCTTG TCCTCCTCCA
1261 GGAAAGCCCC CTTCCTCCTC TGCCCCAGCT CCCAGCCTGG CCTCCTCCAG GCAGGCCCTA
1321 CTCCTCTGCC CCAGCTCCGG CTTTCCCCAT GAGGTTTGTC CCAGGCATGA AGAAAGCATC
1381 CAGGGTGCCA ATGAGTGGGC CTAGGCCAGA GGCCCCTCAG TCCCCAAGGG TACTGTTTTG
1441 GTGGCCTTTC AGAGGGTCAA GGAAGCCCTG CTTGGGGTAG AAGGGGCAGG AGCCCCACAT
1501 GTTGGGGGAG GAAATAAAGT GGAGTGTGCT GTGCTGAAAA AAAAAAAAAA AAAA

```

TAA Stop codon

AATAAA Consensus polyadenylation site (underlined)

FIGURE 10 (continued)

```
1081 CTCGGCAAGG CAGCCTCAGA CCCTAGCTTC CTCCCCAAGA TCAAGGCCTG GAGCAGGTGC
    a s a r q p q t l a s s p r s r p g a g

1141 CCCCgGTGTT GCTCAGGAGC AGTCCTGGCT GGCCGGAGTG TCCACCAAGC CCACAGTCCC
    a p g v a q e q s w l a g v s t k p t v

1201 ATCTTCAGAA GCAGGAATCC AGCCAGTCCC TGTCCAGGGA AGCCCAGCTC TGCCAGGGGG
    p s s e a g i q p v p v q g s p a l p g

1261 CTGTGTACCT AGAAATCATT TCAAGGGGAT GTCCGAAGAT TAAGCCTGTG GCT
    g c v p r n h f k g m s e d
```

78836-FL-R

FIGURE 13

Active site residues are underlined below.

```

WO2002/16566-A2      -----MEGVGGLWPWVLGLLSLPGVILGAPLASSCAGACGTSFPDGLTPEGTQASGDKDI
AX526191             MSCCLVSPVGAPGICVCPCLSGPGVILGAPLASSCAGACGTSFPDGLTPEGTQASGDKDI
INSP005 PREDICTION   -----
INSP005b             -----MEGVGGLWPWVLGLLSLPGVILGAPLASSCAGACGTSFPDGLTPEGTQASGDKDI
INSP005a             -----

WO2002/16566-A2      PAINQGLILEETPESSFLIEGDIIRPSPFRLLSATSSNKWPMGGSGVVEVPFLLSSSKYDEP
AX526191             PAINQGLILEETPESSFLIEGDIIRPSPFRLLSATSSNKWPMGGSGVVEVPFLLSSSKYDEP
INSP005 PREDICTION   -----WPMGGSGVVEVPFLLSSSKYDEP
INSP005b             PAINQGLILEETPESSFLIEGDIIRPSPFRLLSATSSNKWPMGGSGVVEVPFLLSSSKYDEP
INSP005a             -----MGGSGVVEVPFLLSSSKYDEP
                                     *****

WO2002/16566-A2      SRQVILEALAEFERSTCIRFVTYQDQQDFISIIPMYGCFSSVGRSGGMQVVSLAPTCLQK
AX526191             SRQVILEALAEFERSTCIRFVTYQDQQDFISIIPMYGCFSSVGRSGGMQVVSLAPTCLQK
INSP005 PREDICTION   SHQVILEALAEFERSTCIRFVTYQDQQDFISIIPMYGCFSSVGRSGGMQVVSLAPTCLQK
INSP005b             SRQVILEALAEFERSTCIRFVTYQDQQDFISIIPMYGCFSSVGRSGGMQVVSLAPTCLQK
INSP005a             SRQVILEALAEFERSTCIRFVTYQDQQDFISIIPMYGCFSSVGRSGGMQVVSLAPTCLQK
                                     *:*****

WO2002/16566-A2      GRGIVLHELMHVLGFWHEHTRADRDRYIRVNWNEILPGFEINFIKSRSSNMLTPYDYSSV
AX526191             GRGIVLHELMHVLGFWHEHTRADRDRYIRVNWNEILPGFEINFIKSRSSNMLTPYDYSSV
INSP005 PREDICTION   GRGIVLHELMHVLGFWHEHTRADRDRYIRVNWNEILPGFEINFIKSRSSNMLTPYDYSSV
INSP005b             GRGIVLHELMHVLGFWHEHTRADRDRYIRVNWNEILPGFEINFIKSRSSNMLTPYDYSSV
INSP005a             GRGIVLHELMHVLGFWHEHTRADRDRYIRVNWNEILPGFEINFIKSRSSNMLTPYDYSSV
                                     *****:*****

WO2002/16566-A2      MHYGRLAFSRRGLPTITPLWAPSVHIGQRWNLSASDITRVLKLYGCSPSGPRPRGRG---
AX526191             MHYGRLAFSRRGLPTITPLWAPSVHIGQRWNLSASDITRVLKLYGCSPSGPRPRGRGSHA
INSP005 PREDICTION   MHYGRLAFSRRGLPTITPLWAPSVHIGQRWNLSASDITRVLKLYGC-----
INSP005b             MHYGRLAFSRRGLPTITPLWAPSVHIGQRWNLSASDITRVLKLYGCSPSGPRPRGRGSHA
INSP005a             MHYGRLAFSRRGLPTITPLWAPSVHIGQRWNLSASDITRVLKLYGCSPSGPRPRGRGSHA
                                     *****:.. . . .:

```

FIGURE 13 (continued)

```

WO2002/16566-A2 -----EWHG---RKVT
AX526191      HSTGRSPAPASLSLQRLLEALSAESRSPDPSPGSSAGGQVPVPAGPGESPHGWESPALKKLS
INSP005 PREDICTION -----
INSP005b      HSTGRSPAPASLSLQRLLEALSAESRSPDPSPGSSAGGQVPVPAGPGESPHGWESPALKKLS
INSP005a      HSTGRSPAPASLSLQRLLEALSAESRSPDPSPGSSAGGQVPVPAGPGESPHGWESPALKKLS
               :. . :. . : . : :. . :. . :. . :. . :. . :. . :. . :. . :. . :
               :. . :. . :. . :. . :. . :. . :. . :. . :. . :. . :. . :. . :

WO2002/16566-A2 -----
AX526191      AEASARQPQTLASSPRSRPGAGAPGVAQEQSWLAGVSTKPTVPSSEAGIQPVPVQGS
INSP005 PREDICTION -----
INSP005b      AEASARQPQTLASSPRSRPGAGAPGVAQEQSWLAGVSTKPTVPSSEAGIQPVPVQGS
INSP005a      AEASARQPQTLASSPRSRPGAGAPGVAQEQSWLAGVSTKPTVPSSEAGIQPVPVQGS
               :. . :. . :. . :. . :. . :. . :. . :. . :. . :. . :. . :. . :

WO2002/16566-A2 -----
AX526191      PGGCVPRNHFKGMSD
INSP005 PREDICTION -----
INSP005b      PGGCVPRNHFKGMSD
INSP005a      PGGCVPRNHFKGMSD
               .... . . . . :. .

```